# Background

The South Australia Department for Infrastructure and Transport (DIT) has in its possession a wealth of data relating to traffic information collected by way of bluetooth probes, which take count of motor vehicle numbers in a particular area and time, therefore producing a metric for congestion.

This data can be analysed in conjunction with historical real time bus trip updates collected by GTFS-R which, among other metrics, provide the delays in seconds accumulated through a trip’s journey in reference to each stop’s scheduled arrival time.

The aim of the proposed analysis is to detect the extent of the relationship, if any, between the congestion and delay metrics obtained from the AddInsights and the GTFS-R datasets, respectively, for one or more areas and times of interest. The results of the proposed analysis can be used to inform decisions regarding congestion and/or delay mitigating measures.

# Objective

We aim to find the extent of the relationship between road congestion and bus delays for areas and periods of interest.

# Metrics to be Analysed

* Congestion – measurement options are:
  1. Using links data – the quantity of traffic relative to a road’s capacity. A link operating above 1.0 indicates that more vehicles are using the road that it was designed to accommodate under free-flow conditions
  2. Using sites data – the number of unique vehicles identified over a given period per bluetooth probe site, and the average duration spent by a vehicle in that site during that period
* Bus Delay – measured by a stop’s actual arrival time in excess of its scheduled arrival time, in seconds. Idle time and accumulated delays outside the bounds of the area examined will be disregarded.

# Scope

The proposed analysis can be performed through several possible time and location dimensions:

* Time Dimensions
  + Per the morning peak (6am to 10am) and evening peak (3pm to 7pm)
  + Aggregation level(s). For example, 15-minute aggregates, hourly aggregates, …
  + By day of the week
* Location Dimensions
  + Segment of a road
  + A link
  + Bus stop or pair of stops
  + Intersection
  + Towards or away from the city, or both
  + With or without bus lanes
  + North-South/East-West

The time and location dimensions lists are not exhaustive, and we are happy to act through any other dimensions of interest you may identify.

# Road Segment Candidates

The selection criteria we identified for potential road segments to be analysed consist of primarily spatial overlap of bus delay and traffic congestion information, as well as segments with historically high levels of congestions. Possible candidates are:

* South Road
* Main North Road
* Port Wakefield Road
* Glen Osmond Road
* North East Road

We are happy to re-prioritise or expand the list of examined sections to include other areas of interest.

# Data Sources

DIT Transport Analytics AWS datalake. Databases needed:

* addinsight\_prod: to obtain sites and links, and their stats
* gtfsr\_prod: to obtain real-time trip updates that provide stop delay information
* gtfs\_history\_prod: to obtain routes, trips, and stops information

# Proposed Methodology

* Limit analysis to a time period. For example, March 2022, or quarter ending June 2022
  + Disregard weekends and public holidays within that period
* Analysis to be conducted on time-aggregated statistics applied equally to both congestion and delays. For example, 15-minute aggregation resulting in 7am to 7:15am as one aggregate slice, followed by 7:15 to 7:30, 7:30 to 7:45, …
* Limit analysis to segment(s) identified

## Calculating Delays of a Segment

Data obtained from the trip\_updates data set

1. Data preparation:
   1. For each bus trip, remove delays accumulated from stops prior to the segment under examination to isolate the delays to those arising from the concerned segment’s stops only
   2. Remove idle time by recoding negative delays to zero
2. The delay per trip is presented in the dataset as an accumulation of each stop’s incremental delay seconds. Calculate the delay contributed per stop, as an additional measure, to identify bottlenecks and outliers
3. Calculate the average delay per stop per aggregated time across the entire time period to identify possible bottlenecks. For example, the average delay accrued by stop 4916 between 7am and 7:15am across March 2022
4. Calculate the average accumulated delay of all the stops on the segment per aggregated time across the entire time period, to produce the average delay across the entire segment. For example, the average accumulated delay of all the stops between 7am and 7:15am across March 2022

# Calculating Congestion Using Links

Data obtained from the btlinkcongestionstats data set

1. Identify sequence of non-overlapping links that compose the segment
2. Calculate the average congestion per link per aggregated time across the entire time period to identify traffic bottlenecks within a segment
3. Calculate the average congestion of all the links within a segment per aggregated time across the entire time period, to produce the average congestion across the entire segment

The resulting figures are the quantity of traffic relative to a road’s capacity. A link operating above 1.0 indicates that more vehicles are using the road that it was designed to accommodate under free-flow conditions.

## Calculating Congestion Using Sites

Data obtained from the btsitestats data set

Same calculation steps as when using links, however the resulting congestion is measured as the number of vehicles and average time spent by a vehicle in the sites contained within a segment and time aggregate across the time period under analysis.

## Comparison Between Delay and Congestion

Visualisation will be used to examine the relationship between the delay and the congestion per time aggregate. Specific aspects may be highlighted according to the dimensions desired, for example separate visualisations can be constructed according to the direction towards or away from the city, as well the relationship during morning and evening peak hours.

If a relationship is established for a segment, further analysis could be initiated such as:

* Narrowing the time or area scopes to produce outcomes of higher granularity
* Performing the analysis on the same time period from previous years to highlight possible trends
* Examining specific portions of road or bus stops with consistently occurring bottlenecks
* Do certain days exhibit bigger than usual congestions and delays and what might be the cause (for example sporting events)?
* If a bus stop on the segment exhibits primarily no delays on average, does an increase in time taken to reach that stop according to different trips lead to the presence of an incremental delay per that stop? If not, analysis could be performed to determine whether the schedule can be made more efficient by tightening scheduled arrival times

# Example

As an example, the proposed analysis will be implemented on South Road between Ayliffes Road and Richmond Road for the month of March 2022

Map

Description automatically generated

Figure . Segment of South Rd to be analysed highlighted in blue

Below is the comparison between delays and congestion using link stats, separated by direction of travel. The highlighted areas are the morning and evening peaks. The data were aggregated using 15-minute intervals

Chart, bar chart, histogram

Description automatically generated

Figure . Delays vs congestion using link stats to/from the city with peaks highlighted

Below is the comparison using sites stats

Chart, line chart

Description automatically generated

Figure . Delays vs congestion using site stats with peaks highlighted

The figures show that delays increase along with congestion during the evening peak when leaving the city. Delays also increase in the evening towards the city.

# Estimated Completion Date

October 7, 2022